Amendments to the Specification:

Please replace paragraph [0024] beginning on page 11, with the following amended paragraph:

[0024] FIG. 4 is a block diagram of an example audio search module 400 that performs robust audio search. Audio search module 400 comprises a feature extractor 410, a modeling mechanism 420, and a decision maker 430. Feature extractor 410 may receive an input audio stream (e.g., a target audio clip, a substream of a large audio stream, etc.) and extract acoustic features from the input audio stream. When the input audio stream is an audio stream to be searched for the target audio clip, the feature extractor may apply sliding window on the audio stream to partition it into multiple overlapped segments. The window has the same length as the target audio clip. Each segment of the input audio stream (the target audio stream has only one segment) is further separated into frames. Each frame may have the same length and may overlap with its neighboring frames. For example, in one embodiment, a frame may be 20 milliseconds in length with the overlap between frames being 10 milliseconds. A feature vector may be extracted for each frame, which may include such features as Fourier coefficients, Mel-Frequency cepstral coefficients, spectral flatness, and means, variances, and other derivatives thereof. Feature vectors from all of the frames in an audio segment form a feature vector sequence.